J Stat Phys (2013) 152:1170–1193
DOI 10.1007/s10955-013-0801-1

Existence of Nonequilibrium Steady State for a Simple Model of Heat Conduction

Yao Li · Lai-Sang Young

Received: 15 February 2013 / Accepted: 2 July 2013 / Published online: 24 July 2013

Abstract This paper contains rigorous results for a simple stochastic model of heat conduction similar to the KMP (Knipnis–Marchiori–Presutti) model but with possibly energy-dependent interaction rates. We prove the existence and uniqueness of nonequilibrium steady states, their relation to Lebesgue measure, and exponential convergence to steady states from suitable initial conditions.

Keywords Nonequilibrium steady states · Energy exchange · Markov chains

1 Introduction

A number of mathematical models of 1-D heat conduction have been introduced in recent years. Some of these models are defined by purely deterministic dynamics (1, 5, 7, 8, 11, 12, 14, 15, 21, 22, 24, 26, 27, 33–35), some are deterministic systems perturbed by noise ([2, 3, 25]), and others are stochastic models ([10, 12, 16, 20, 30]). These setups provide useful frameworks for studying nonequilibrium phenomena such as conductivity, energy and density profiles, Fourier Law, fluctuations and large deviations, see e.g., [4, 9, 23]. From a mathematical point of view, among the most basic questions are the existence and uniqueness of nonequilibrium steady states when the chain is connected to unequal heat baths, and the rates of convergence to steady states. Fundamental as they are, these questions are very difficult for Hamiltonian systems. The only situation for which there is a complete understanding is that of chains of anharmonic oscillators, a summary of the results for which is given in [31]. Stochastic models are generally simpler; still, answers have been elusive, especially with regard to questions of existence of invariant measures.

LSY was supported in part by NSF Grant DMS-1101594.

Y. Li · L.-S. Young (✉)
Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA
e-mail: lsy@cims.nyu.edu

Y. Li
e-mail: yaoli@cims.nyu.edu

The purpose of this paper is to illustrate, using a simple stochastic model, how these questions can sometimes be settled using certain known techniques. Inspired by the KMP-model [20], we consider here a model consisting of a chain of N sites each of which is described by a single quantity, namely its energy. When the clock between two adjacent sites rings, the energies of the two sites are pooled together and randomly repartitioned. A similar operation is performed between the end sites and the baths: *i.i.d.* sequences of energies that are exponentially distributed are drawn from the baths and mixed with energies from the end sites. Unlike the KMP-model, where all the clocks have rate 1, we permit energy-dependent rates of interactions. This may be both more realistic and more natural for particle systems: one can speed up the dynamics by speeding up individual particles, and in a chain that is out of equilibrium, particles are likely to move at different speeds in different parts of the chain. For a class of models with these properties (see Sect. 2.1 for a precise description), we prove the existence and uniqueness of nonequilibrium steady states, their absolute continuity with respect to Lebesgue measure, and the exponential convergence to steady states starting from suitable initial conditions. Precise results are stated in Sect. 3. Generalized KMP processes have also been considered recently in the physics literature; see e.g. [29].

Conceptually, the following are the main issues: For the existence of invariant measures, one needs to control the amount of energy that flows into and out of a chain. This can be problematic because interaction with heat baths is determined by the dynamics near the ends of the chain, and they may not adequately reflect what goes on in the interior. Pathologies such as the hoarding of energy, which leads to a chain's heating up, or the reverse, leading to its freezing [13], can in principle happen; it is one thing to say they should not happen for models that are "sufficiently physical", another to mathematically rule them out. As for uniqueness of the invariant measure, and its absolute continuity with respect to Lebesgue measure, the main difficulty is that in systems with nearest-neighbor interactions, such as what we have here, transition probabilities corresponding to individual interactions are highly degenerate, as they involve only two out of the N variables (or sets of variables) where N is the length of the chain.

With regard to technical proofs, for the type of problems considered here one generally uses either operator based or probabilistic arguments. In a spectral approach, the existence of invariant measures is obtained by solving an eigenvalue problem, and exponential convergence corresponds to a spectral gap. For the nonequilibrium results on anharmonic chains, the existence and uniqueness of invariant measures was first proved in [14] using these techniques (though the authors did not prove a spectral gap). In more probabilistic approaches, Lyapunov functions are often used to ensure tightness, and coupling (see e.g. [28]) is by now a standard tool for studying the speed of convergence. For anharmonic chains, these ideas were used in [27] to show exponential convergence to steady state; this paper also contains a second proof of existence. We have elected to go the probabilistic route, which avoids technical choices of functions spaces. While the present model is simple enough that both methods are likely workable, we think that the Lyapunov-function-coupling approach, which is "softer", may be easier to apply to larger classes of nonequilibrium stochastic models. Finally, the systems considered in [8, 14, 27] are defined by SDEs, for which Hörmander's conditions can be used to give hypoellipticity. That in turn implies that every invariant measure has a density supported on the entire space, from which the uniqueness of invariant measure follows. These ideas are used in the papers cited. Our model is not described by an SDE, and we know of no such ready-made tools applicable to Markov jump processes of the type considered here; hence we have to prove these results "by hand".

The organization of this paper is straightforward: Model description is given in Sect. 2, followed by statements of results in Sect. 3. Background material is reviewed in Sect. 4;

main ingredients of the proofs are contained in Sects. 5 and 6, and the proofs are completed in Sect. 7.

2 Model Description

2.1 Heat Transport Model of KMP Type

We consider in this paper a heat transport model having a setup similar to that in [20] but with a more general, possibly energy-dependent, interaction rate.

Let $N \geq 1$ be a fixed integer. We consider N linearly ordered *sites* labeled $1, 2, \dots, N$, each storing a finite amount of energy $x_i > 0$, so that the *states* of the model are represented by vectors $\mathbf{x} = (x_1, x_2, \dots, x_N) \in \mathbb{R}_+^N$. Energy exchanges take place between adjacent sites, and the two ends of the chain are coupled to two *heat baths*, which we think of as located at sites 0 and $N+1$. The temperatures of the heat baths are set at T_L and T_R. Without loss of generality, we assume $0 < T_L, T_R < 1$.

More precisely, the system is equipped with independent exponential *clocks*, $c_0, c_1, \dots, c_N$. For $0 < i < N$, c_i determines the times of energy exchange between sites i and $i+1$: When it goes off, x_i and x_{i+1} are updated instantaneously to $\hat{x}_i$ and $\hat{x}_{i+1}$ where

$$\hat{x}_i = p(x_i + x_{i+1}) \quad \text{and} \quad \hat{x}_{i+1} = (1-p)(x_i + x_{i+1}),$$

p being a uniformly distributed random variable on [0, 1] that is independent of everything else. Similarly, c_0 and c_N signal the times of energy exchange between the two baths and the sites closest to each. When c_0 rings, x_1 is updated to $\hat{x}_1 = p(x_1 + y)$ where y is drawn randomly from an exponential distribution of mean T_L; the choice of y is independent of history or anything else. Likewise, when c_N rings, x_N is updated to $\hat{x}_N = p(x_N + z)$ where z is drawn randomly and independently from an exponential distribution of mean T_R.

In the original KMP model introduced in [20], all $N+1$ clocks ring at rate 1. Here we relax this condition to permit energy-dependent rates of interaction. We assume that η_i, the rate of clock c_i, varies with time and is a function of the energies of the associated sites at that moment, i.e., for $0 < i < N$, $\eta_i = f(x_i, x_{i+1})$ for a function f for which certain technical conditions will be imposed. For the rates of interaction with baths, we assume $\eta_0 = f(T_L, x_1)$ and $\eta_N = f(x_N, T_R)$.

We assume the following conditions are satisfied by the function $f : \mathbb{R}_+^2 \to (0, \infty)$:

Assumptions (H)

(1) There is a nondecreasing function $h(\cdot)$ and a constant C such that

$$h\big(\max\{x, y\}\big) \leq f(x, y) \leq Ch\big(\max\{x, y\}\big);$$

(2) $h(\cdot)$ has sublinear growth rate, i.e., $h(\alpha z) \leq \alpha h(z)$ for all $\alpha \geq 1$.

Notice that by setting $h \equiv 1$ and $C = 1$, we recover the case of $\eta_i \equiv 1$. Our motivation for relaxing to energy-dependent rates of interaction comes from several examples of Hamiltonian (particle) systems (see e.g. [6, 12, 15, 34]), in which these rates are clearly affected by the energies of the particles in the relevant sites though the situation is nowhere near as simple as that described above.

2.2 Mathematical Framework

The model described in Sect. 2.1 is a continuous-time Markov chain on $\mathbb{R}_+^N$. The state of the system at time t is denoted by

$$\mathbf{x}(t) = \big(x_1(t), x_2(t), \ldots, x_N(t)\big) \in \mathbb{R}_+^N := (0, \infty)^N.$$

Let $\mathcal{B} = \mathcal{B}(\mathbb{R}_+^N)$ be the Borel σ-algebra on $\mathbb{R}_+^N$. Throughout this paper, we will refer to this Markov chain as $\Phi_t, t \geq 0$, and its *Markov transition kernel* as $P^t(x, A)$. This implies in particular that for $t \geq 0$, $P^t(\mathbf{x}, \cdot)$ is a probability measure on $(\mathbb{R}_+^N, \mathcal{B})$ for every $\mathbf{x} \in \mathbb{R}_+^N$, and $P^t(\cdot, A)$ is a measurable function for every $A \in \mathcal{B}$.

Formally, Φ_t is defined by the infinitesimal generator $\mathcal{G}$ acting on bounded measurable functions ξ on $\mathbb{R}_+^N$ as follows:

$$\begin{aligned}
&(\mathcal{G}\xi)(\mathbf{x}) \\
&\quad = \sum_{i=1}^{N-1} f(x_i, x_{i+1}) \int_0^1 \big[\xi\big(x_1, \ldots, p(x_i + x_{i+1}), (1-p)(x_i + x_{i+1}), \ldots, x_N\big) - \xi(\mathbf{x})\big] \mathrm{d}p \\
&\qquad + f(T_L, x_1) \int_0^\infty \int_0^1 \big[\xi\big(p(x_1 + y), x_2, \ldots, x_N\big) - \xi(\mathbf{x})\big] \beta_L e^{-\beta_L y} \mathrm{d}p \mathrm{d}y \\
&\qquad + f(x_N, T_R) \int_0^\infty \int_0^1 \big[\xi\big(x_1, x_2, \ldots, p(x_N + y)\big) - \xi(\mathbf{x})\big] \beta_R e^{-\beta_R y} \mathrm{d}p \mathrm{d}y,
\end{aligned}$$

where $\beta_L = (T_L)^{-1}$ and $\beta_R = (T_R)^{-1}$.

Given $\Delta > 0$, the *time-*Δ *sample chain* of Φ_t is the discrete-time Markov chain Φ_n^Δ, $n = 0, 1, 2, \ldots$, whose transition probabilities are given by $P^\Delta(\mathbf{x}, \cdot)$. For $\Delta = 1$, we omit the superscript Δ, and write simply Φ_n and $P(\mathbf{x}, \cdot)$.

The left operator associated with the Markov transition kernel $P^t(\mathbf{x}, A)$ acting on the set of probability measures on $(\mathbb{R}_+^N, \mathcal{B})$ is given by

$$\big(\mu P^t\big)(A) = \int_{\mathbb{R}_+^N} P^t(\mathbf{x}, A)\mu(\mathrm{d}\mathbf{x}); \tag{2.1}$$

the right operator acting on the set of bounded measurable functions on $(\mathbb{R}_+^N, \mathcal{B})$ is given by

$$\big(P^t\xi\big)(\mathbf{x}) = \int_{\mathbb{R}_+^N} \xi(y) P^t(\mathbf{x}, \mathrm{d}y). \tag{2.2}$$

Left and right operators associated with the time-1 chain Φ_n are defined similarly.

A probability measure μ on $(\mathbb{R}_+^N, \mathcal{B})$ is said to be *invariant* under the Markov chain Φ_t (resp. Φ_n) if $\mu P^t = \mu$ for all $t \geq 0$ (resp. $\mu P = \mu$).

2.3 Justification of Rate Function

While it is not unnatural to assume that higher energy leads to faster clock rates, it is not true that the rates of interaction between adjacent sites carrying energies x and y can always be represented as a function of $\max\{x, y\}$. We provide here an example (see Fig. 1) that lends some justification to the form of $f(x, y)$ in Assumptions (H).

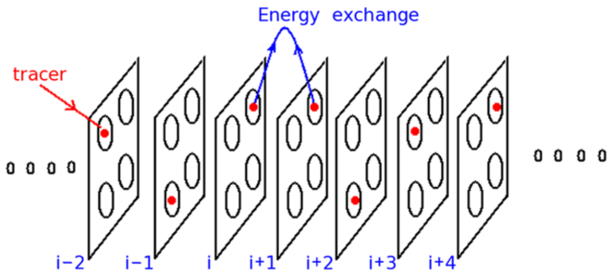


Fig. 1 A heat transport model with m internal states in each site

We assume that (i) within each site there are m internal states, labeled $l_1, l_2, \dots, l_m$; (ii) in site i a particle carrying energy x_i hops from one internal state to another, its destination chosen randomly and independently of anything else, and the times between hops are exponentially distributed with mean $\sqrt{x}_i$; (iii) energy exchange between sites i and $i+1$ occurs at the first instant when the particles in the two sites find themselves in the same internal state; and (iv) after an exchange, the process continues with the particles hopping at rates determined by their new energies.

For argument sake, we assume that at time 0, the particles in sites i and $i+1$ have internal states $l_{\sigma(i)} \neq l_{\sigma(i+1)}$. We let $v_i = \sqrt{x_i}$, $v_{i+1} = \sqrt{x_{i+1}}$, $v_{\max} = \max\{v_i, v_{i+1}\}$, and assume $v_i \geq v_{i+1}$. In a small time interval $[0, \varepsilon]$, the probability that energy is exchanged is larger than the probability that the particle in site i hops to $l_{\sigma(i+1)}$ while the particle in site $i+1$ does not move. Hence

$$f(x, y) \geq \lim_{\varepsilon \to 0} \frac{\frac{1}{m}(1 - e^{-v_i \varepsilon}) e^{-v_{i+1}\varepsilon}}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{v_i}{m} e^{-v_{i+1}\varepsilon} = \frac{1}{m} v_{\max}.$$

On the other hand the probability that no energy is exchanged during the time period $[0, \varepsilon]$ is larger than the probability that neither particle changed their internal states. Hence

$$f(x, y) \leq \lim_{\varepsilon \to 0} \frac{1 - e^{-v_i \varepsilon} e^{-v_{i+1}\varepsilon}}{\varepsilon} \leq 2 v_{\max}.$$

Hence the assumptions in Assumptions (H) are satisfied with $h(z) = \frac{1}{m}\sqrt{z}$ and $C = 2$.

3 Statement of Main Results

We begin by defining, in a general setting, a few terms that appear in the statement of our results. Let $(X, \mathcal{A})$ be a measurable space, and let $W : X \to [1, \infty)$ be a measurable function on X. We define the *W-weighted supremum norm* of a measurable function $\xi : X \to \mathbb{R}$ to be

$$\|\xi\|_W = \sup_{x \in X} \frac{|\xi(x)|}{W(x)},$$

so that for $W \equiv 1$, $\|\xi\| := \|\xi\|_1 = \sup_{x \in X} |\xi(x)|$ is the usual supremum norm of ξ. We also define the *W-weighted variation norm* of a signed measure μ on $(X, \mathcal{A})$ to be

$$\|\mu\|_W = \int_X W(x) |\mu|(\mathrm{d}x),$$

and let $L_W(X, \mathcal{A})$, or simply $L_W(X)$, denote the set of probability measures μ on $(X, \mathcal{A})$ with $\mu(W) < \infty$.

Given a Markov chain Ψ_t on X with Markovian transition kernel $\hat{P}^t$, the *time correlation function* of observables ξ and ζ with respect to a probability measure μ on $(X, \mathcal{A})$ is defined to be

$$C^{\mu}_{\xi,\zeta}(t) = \int_X \big(\hat{P}^t\zeta\big)(x)\xi(x)\mu(\mathrm{d}x) - \int_X \big(\hat{P}^t\zeta\big)(x)\mu(\mathrm{d}x)\int_X \xi(x)\mu(\mathrm{d}x).$$

The theorems below apply to the Markov chain Φ_t defined in Sect. 2.2; system size N and bath temperatures $T_L, T_R < 1$ are to be regarded as fixed. The interaction rates $f(x, y)$ of the chain are assumed to satisfy the conditions in Assumptions (H). Let $V : \mathbb{R}^N_+ \to [1, \infty)$ be given by

$$V(\mathbf{x}) = 1 + V_0(\mathbf{x}) \quad \text{where } V_0(\mathbf{x}) = \sum_{i=1}^{N} x_i \tag{3.1}$$

is the total energy of the state $\mathbf{x}$. Lebesgue measure on $\mathbb{R}^N_+$ is denoted by λ.

Theorem 1 *There exists a unique $\pi \in L_V(\mathbb{R}^N_+, \mathcal{B})$ that is invariant under Φ_t.*

Theorem 1 tells us where an invariant measure can be found; it does not guarantee that Φ_t has no other invariant measure (outside of $L_V(\mathbb{R}^N_+, \mathcal{B})$), nor does it tell us if π has a density. These two issues are not unrelated; we record the results in Theorem 2.

Theorem 2

(a) *The measure π is absolutely continuous with respect to λ, with $\frac{d\pi}{d\lambda} > 0$ λ-a.e.*
(b) *It is the only Φ_t-invariant probability measure.*

Having established the existence and uniqueness of nonequilibrium steady state, we proceed to describe its statistical properties. The ergodicity of (Φ_t, π) is an immediate consequence of Theorem 2(b). Our next results concern the speed of mixing.

Theorem 3 *There exist constants $c > 0$ and $0 < \rho < 1$ such that*

(a) $\|P^t(\mathbf{x}, \cdot) - \pi\|_V \le cV(\mathbf{x})\rho^t$ *for every* $\mathbf{x} \in \mathbb{R}^N_+$;
(b) *more generally, $\mu \in L_V(\mathbb{R}^N_+, \mathcal{B})$ implies $\mu P^t \in L_V(\mathbb{R}^N_+, \mathcal{B})$, and*

$$\big\|\mu_1 P^t - \mu_2 P^t\big\|_V \le c\|\mu_1 - \mu_2\|_V \rho^t \quad \textit{for all } \mu_1, \mu_2 \in L_V\big(\mathbb{R}^N_+\big).$$

Corollary 4 *The Markov chain Φ_t has exponential decay of time correlations with respect to the following measure and function classes: Let $\mu \in L_V(\mathbb{R}^N_+)$, and let ξ and ζ be measurable functions on $(\mathbb{R}^N_+, \mathcal{B})$ with $\|\xi\| < \infty$ and $\|\zeta\|_V < \infty$. Then*

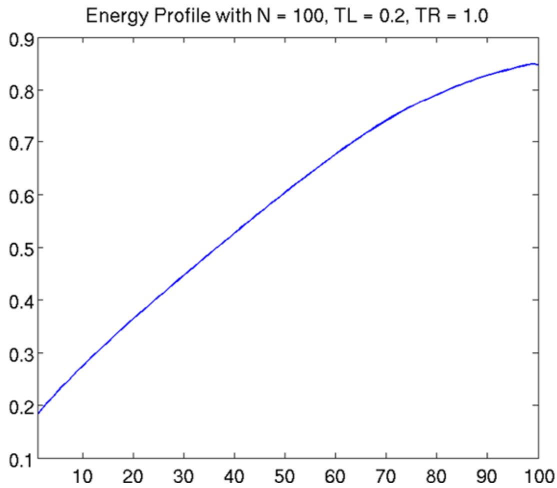
$$\big|C^{\mu}_{\xi,\zeta}(t)\big| \le \|\xi\|\|\zeta\|_V\|\mu\|_V \cdot 2c\rho^t$$

where c and ρ are as in Theorem 3.

Remarks

(1) Tightness is needed to establish the existence of invariant measures, due to the noncompactness of the domain $\mathbb{R}^N_+$. For the case considered in [20], i.e., where all the clocks have rate 1, this follows from a simple coupling argument: Observe first that when

Fig. 2 Energy profile of a chain with energy dependent interactions. Here, $N = 100$, $T_L = 0.2$, $T_R = 1.0$ and $f(x, y) = \sqrt{\max\{x, y\}}$



$T_L = T_R = T$, the product measure $\Pi_{i=1}^N \beta e^{-\beta x_i}$, where $\beta = T^{-1}$, is invariant. Given a system with $T_L < T_R$, one can couple it to the system with $\bar{T}_L = \bar{T}_R = T_R$ as follows: Notice that each sample path is determined by $(t_k, i_k, p_k; e_k^L, e_k^R)_{k=1,2,\ldots}$ where $t_1 < t_2 < \cdots$ are the times at which mixing occurs, x_{i_k} and x_{i_k+1} are mixed at time t_k with proportions p_k and $1 - p_k$, and e_k^L and e_k^R are the kth energies drawn from the two baths. We can couple the system with baths (T_L, T_R) to the one with baths $(\bar{T}_L, \bar{T}_R)$ sample path by sample path, i.e., start from identical initial distributions, and use for both the same sequence $(t_k, i_k, p_k; *, e_k^R)$ except in the slot marked $*$, where we require each energy drawn from the T_L-bath be less than the corresponding energy drawn from the $\bar{T}_L$-bath. This way, every x_i in the (T_L, T_R)-system is smaller than the corresponding entry in the $(\bar{T}_L, \bar{T}_R)$-system at all times, guaranteeing that its invariant measure is dominated by the invariant measure of the $(\bar{T}_L, \bar{T}_R)$-system.

Notice that this argument fails when the rates of interactions are energy-dependent, for when the clocks have variable rates, there is no natural coupling of sample paths that will preserve the *order* of the interactions, and without this order preservation, the energies $\{x_i\}$ of the two systems are not directly comparable.

(2) Even when the setups are similar, nonequilibrium steady states can be qualitatively quite different for energy-dependent and energy-independent interactions. For example, for a fixed N, let E_i be the mean of the ith marginal distribution of π, i.e., $E_i = \mathbb{E}_\pi[x_i]$. We rescale this information onto $(0, 1)$ by defining $E^N(x) = E_i$ for $x \in (\frac{i-1}{N}, \frac{i}{N})$ for $i = 1, 2, \ldots, N$. In the case where $\eta_i \equiv 1$, E^N tends to a linear profile $E^\infty(x) = T_L + (T_R - T_L)x$; this is proved in [20]. In the case where higher energy leads to faster clocks, we believe E^∞ can be strictly concave for $T_R > T_L$, as the energy at site i is "more influenced" by the distribution in site $i + 1$ than the one in site $i - 1$. A concrete example is given in Fig. 2.

4 Method of Proof

4.1 Some Known Results

We review here some known results on discrete-time Markov chains on general spaces that we will use. These results go back to Harris [18], who showed that unique ergodicity follows from the existence of "small sets" that are visited infinitely often, extending both the picture for countable state Markov chains certain ideas of Doeblin to "unbounded" state spaces. This "small set" condition is often established by finding a Lyapunov function whose level sets have certain properties [19, 28]; a detailed exposition is contained in [28]. We will, however, follow [17], which contains a more direct proof of these results and has a formulation well suited for our purposes.

The setting is as follows: Let $(X, \mathcal{A})$ be a measurable space, and let $\Psi_n, n = 0, 1, 2, \dots,$ be an arbitrary Markov chain on X with Markov transition kernel $\hat{P}(x, \cdot)$. As before, we write

$$(\hat{P}\xi)(x) = \int_X \hat{P}(x, \mathrm{d}y)\xi(y) \quad \text{and} \quad (\mu\hat{P})(A) = \int_X \hat{P}(x, A)\mu(\mathrm{d}x).$$

Here $\xi : X \to \mathbb{R}$ is a bounded measurable function and μ is a probability measure on $(X, \mathcal{A})$. Of interest are the following two conditions:

Assumption (A1) There exists a measurable function $W_0 : X \to [0, \infty)$ and constants $0 \le K < +\infty$, $\gamma \in (0, 1)$ such that

$$(\hat{P}W_0)(x) - W_0(x) \le -\gamma W_0(x) + K \quad \text{for all } x \in X.$$

Assumption (A2) Let $D = \{W_0 \le 2K/\gamma\}$. Then there exist $\alpha \in (0, 1)$ and a probability measures ν on $(X, \mathcal{A})$ such that

$$\inf_{x\in D} \hat{P}(x, \cdot) \ge \alpha\nu(\cdot).$$

These assumptions carry the following intuitive meaning: Assumption (A1) asserts the existence of a Lyapunov type function; the W_0-value of a state is, on average, decreased unless it is already below a certain value. This is clearly conducive to tightness. Assumption (A2) asserts a Doeblin type condition for a set of the form $\{W_0 \le R\}$ for R large enough. Starting from two initial distributions, these conditions say that the tendency of the chain is to drive them – uniformly but in the sense of expectations only – into a set from which a fraction of the measures can be coupled, setting the stage for exponential mixing.

In what follows, we let $W = W_0 + 1$, and let $\|\xi\|_W$, $\|\mu\|_W$ and $L_W(X)$ be as defined in the beginning of Sect. 3. It follows easily from Assumption (A1) that for $\mu \in L_W(X)$, $\mu\hat{P}$ is also in $L_W(X)$, because $(\mu\hat{P})(W) \le (1-\gamma)\mu(W) + K$.

We recall the following results from [17]:

Theorem 4.1 (Theorems 1.2 and 1.3 in [17]) *Assume Assumptions (A1) and (A2) hold for the Markov chain Ψ_n. Then Ψ_n has a unique invariant probability measure μ_* in $L_W(X)$. Moreover, there exist $\bar{\alpha} \in (0, 1)$ and $C_1, C_2 > 0$ such that*

(a) *for all $\mu_1, \mu_2 \in L_W(X)$,*

$$\left\|\mu_1\hat{P}^n - \mu_2\hat{P}^n\right\|_W \le C_1\bar{\alpha}^n\|\mu_1 - \mu_2\|_W;$$

(b) *for every measurable function ξ with $\|\xi\|_W < \infty$,*

$$\left\|\hat{P}^n\xi - \mu_*(\xi)\right\|_W \le C_2\bar{\alpha}^n\left\|\xi - \mu_*(\xi)\right\|_W.$$

Remarks

(1) In [17], a family of norms parameterized by $\beta > 0$ is introduced using $W_\beta = 1 + \beta W_0$ in the place of W. The authors of [17] showed that for suitable choices of β, the action of $\hat{P}$ is a contraction on $L_{W_\beta}(X)$, i.e., for $\mu_1, \mu_2 \in L_{W_\beta}(X)$,

$$\|\mu_1 \hat{P} - \mu_2 \hat{P}\|_{W_\beta} \le \bar{\alpha} \|\mu_1 - \mu_2\|_{W_\beta}.$$

The results above are deduced from this together the facts that (i) for any two β, the corresponding norms are equivalent, and (ii) equipped with these norms, the spaces $L_{W_\beta}(X)$ are complete.

(2) This remark pertains to the constants $\bar{\alpha}$, C_1 and C_2 in Theorem 4.1. It is shown in [17] that if Assumption (A2) holds on $D_R = \{W \le R\}$ for $R \ge 2K/\gamma$ where γ and K are the constants in Assumption (A1), then for any $\alpha_* \in (0, \alpha)$ and any $\gamma_* \in (1 - \gamma + 2K/R, 1)$, one can choose $\beta = \alpha_*/K$ and obtain a contraction rate

$$\bar{\alpha} = \max\{1 - (\alpha - \alpha_0), (2 + R\beta\gamma_*)/(2 + R\beta)\}$$

with respect to the $\|\cdot\|_{W_\beta}$-norm (see item (1) above). The constants C_1 and C_2 come from the equivalence of the norms $\|\cdot\|_W$ and $\|\cdot\|_{W_\beta}$.

4.2 Outline of Proofs of Our Results

To prove Theorems 1 and 3 stated in Sect. 3, we will apply the results in Sect. 4.1 to the time-1 sample chain Φ_n of Φ_t. The proof is divided into the following three steps:

(i) The first step is to establish Assumption (A1) with $W_0 = V_0$, the total energy function. This consists of proving the following: (a) In the time interval $[0, 1)$, the expected energy gain under Φ_t is uniformly bounded independent of initial condition, and (b) starting from any initial distribution, there is a positive probability (bounded away from zero) that a fraction of the initial energy will be released into the heat baths before time 1.

(ii) The second step is to establish Assumption (A2). We will show in our case that the set D in Assumption (A2) can be taken to be $\{V_0 \le M\}$ for any $M \in (1, \infty)$.

(iii) With the successful completion of these two steps, we will be in a position to appeal to Theorem 4.1 to obtain results analogous to Theorems 1 and 3 for the *time-1 chain* Φ_n. In the last step, we pass these results to the original continuous-time Markov chain Φ_t. Corollary 4 is easily deduced from Theorem 3.

The proof of Theorem 2 is related to Assumption (A2) but involves some additional ideas. We will show that *any* invariant probability measure of Φ_t is necessarily absolutely continuous with respect to Lebesgue measure, with a strictly positive density almost everywhere.

5 Expected Energy Gains

To prove tightness, or that the chain does not heat up over time, we need to control the amount of energy that flows into the chain. More precisely, given an initial condition $\mathbf{x}(0)$ and a number $\tau > 0$, we define

$$G_{L,[0,\tau)}(\mathbf{x}(0)) := \mathbb{E}_{\mathbf{x}(0)}\left[\sum_{t_\alpha} \max\{x_1(t_\alpha^+) - x_1(t_\alpha), 0\} \mathbf{1}_{\{0<t_\alpha<\tau\}}\right]$$

where $t_\alpha > 0$ are the times at which c_0 rings, $x_1(t_\alpha)$ is the energy at site 1 before the exchange and $x_1(t_\alpha^+)$ immediately after the exchange. Similarly, define $G_{R,[0,\tau)}(\mathbf{x}(0))$ replacing c_0 and x_1 by c_N and x_N respectively, and let $G := G_L + G_R$. Notice that $G \geq 0$, as it counts only the net flow of energy from the bath into the chain in each interaction. If the net flow is from the chain to the bath, a "0" is registered. We will refer to $G_{[0,1)}(\mathbf{x}(0))$ as the expected energy gain by the chain per unit time starting from initial condition $\mathbf{x}(0)$.

The goal of this section is to prove

Proposition 5.1 *There exists a constant B depending on $Ch(1)$ (and the assumption that T_L, $T_R < 1$) such that $G_{[0,1)}(\mathbf{x}(0)) \leq B$ for all $\mathbf{x}(0) \in \mathbb{R}_+^N$.*

For the definition of "$Ch(1)$", see Assumptions (H) in Sect. 2.1. The rationale for this result is that when x_1 is small, clock c_0 should not ring too often, and when it is large, the net flow of energy is more likely from the chain to the bath when c_0 rings, and the same is true for x_N and c_N. Below we make these ideas mathematically rigorous.

5.1 Main Computation

Observe that before the first energy exchange with a bath, the rate η_i of clock c_i is less than or equal to $\max\{1, V_0(\mathbf{x}(0))Ch(1)\}$ for all i. We check this for c_0: If $\max\{T_L, x_1\} \leq 1$, then $f(T_L, x_1) \leq Ch(\max\{T_L, x_1\}) \leq Ch(1)$ by the monotonicity of h. If $\max\{T_L, x_1\} > 1$, then $\max\{T_L, x_1\} = x_1 > 1$, and $f(T_L, x_1) \leq x_1 Ch(1) \leq V_0(\mathbf{x}(0))Ch(1)$ by condition (2) in Assumptions (H). The other cases are similar.

The rest of this section is devoted to the proof of the following lemma:

Lemma 5.2 *There exists a constant B such that for every $\mathbf{x}(0) \in \mathbb{R}_+^N$, if $\tau \leq Ch(1)e^{-2}M^{-2}$ for some $M \geq \max\{2, V_0(\mathbf{x}(0))\}Ch(1)$, then $G_{[0,\tau)}(\mathbf{x}(0)) \leq B\tau$.*

We note that both M and τ depend on $\mathbf{x}(0)$, and for as long as $V_0 \leq V_0(\mathbf{x}(0))$, we have $\eta_i \leq M$ (the "2" in its definition is for later convenience), and $M\tau \leq \frac{1}{2}e^{-2}$.

Proof We fix $\mathbf{x}(0)$, and consider separately the following mutually exclusive cases. It will be assumed implicitly throughout that all discussions pertain to the interval $[0, \tau)$. For $k = 0, 1, 2, \ldots, \infty$, let

$$E_k = \{c_0 \text{ and } c_N \text{ together ring exactly } k \text{ times}\}.$$

The contributions to $G_{[0,\tau)}(\mathbf{x}(0))$ by the various events are analyzed as follows:

Event E_0: Clearly, there is no energy gain in this case.

Event E_1: We let $E_1 = E_1^0 \cup E_1^N$ corresponding to the cases where c_0 or c_N rings. Since the two cases are treated in an identical fashion, we consider c_0 and drop the superscript 0. We further subdivide E_1 (meaning E_1^0) into the following two cases:

$$E_{1a} := \{c_0 \text{ rings exactly once}, c_1 \text{ does not ring before } c_0; c_N \text{ does not ring}\};$$

$$E_{1b} := \{c_0 \text{ rings exactly once}, c_1 \text{ rings before } c_0; c_N \text{ does not ring}\}.$$

For notational simplicity, let us write $\beta = \beta_L$ and $x_1 = x_1(0)$. Since $\mathbb{P}[E_{1a}] \leq \mathbb{P}[c_0 \text{ rings exactly once} \mid c_1 \text{ does not ring before } c_0, \text{ and } c_N \text{ does not ring}]$, we have

$$\begin{aligned}\mathbb{P}[E_{1a}] &\le \int_0^\infty\int_0^1\int_0^\tau e^{-f(x_1,T_L)s} f(x_1,T_L)e^{-f(p(x_1+y),T_L)(\tau-s)}\beta e^{-\beta y}\mathrm{d}s\mathrm{d}p\mathrm{d}y\\
&\le f(x_1,T_L)\int_0^\infty\int_0^1\int_0^\tau e^{f(p(x_1+y),T_L)s}\beta e^{-\beta y}\mathrm{d}s\mathrm{d}p\mathrm{d}y\\
&\le f(x_1,T_L)\tau\int_0^\infty e^{(x_1+y+1)Ch(1)\tau}\beta e^{-\beta y}\mathrm{d}y\\
&= f(x_1,T_L)\tau\cdot\frac{\beta}{\beta-Ch(1)\tau}\cdot e^{(x_1+1)Ch(1)\tau}\\
&< 2f(x_1,T_L)\tau.\end{aligned}$$

To bound the last inequality, we have used $Ch(1)\tau \le \frac{1}{4}e^{-2}$ and $x_1Ch(1) \le M$.

It follows that the expected energy gain associated with E_{1a} is given by

$$\begin{aligned}\mathbb{P}[E_{1a}]\cdot\mathbb{E}\big[G_{[0,\tau)}\big(\mathbf{x}(0)\big)\mid E_{1a}\big] &= \mathbb{P}[E_{1a}]\int_0^\infty\int_0^1 \max\big\{p(x_1+y)-x_1,0\big\}\beta e^{-\beta y}\mathrm{d}p\mathrm{d}y\\
&< 2\tau f(T_L,x_1)\int_0^\infty \frac{y^2}{2(x_1+y)}\beta e^{-\beta y}\mathrm{d}y.\end{aligned}$$

Since

$$f(T_L,x_1)\le Ch\big(\max\{x_1,T_L\}\big) < Ch(x_1+1)\le (x_1+1)Ch(1)$$

and

$$\frac{y^2}{x_1+y} < \frac{1+y^2}{x_1+1}\quad \text{for all } x_1, y>0,$$

it follows that

$$\begin{aligned}\mathbb{P}[E_{1a}]\cdot\mathbb{E}\big[G_{[0,\tau)}\big(\mathbf{x}(0)\big)\mid E_{1a}\big] &< Ch(1)\tau\int_0^\infty \big(1+y^2\big)\beta e^{-\beta y}\mathrm{d}y\\
&= Ch(1)\big(1+2\beta^{-2}\big)\tau < 3Ch(1)\tau.\end{aligned}$$

Next, we claim

$$\mathbb{P}[E_{1b}]\le\big(1-e^{-M\tau}\big)2M\tau < 2M^2\tau^2.$$

For as long as $V_0 = V_0(\mathbf{x}(0))$ up to the time of the ring, the probability of c_i ringing at least once is the quantity in brackets, and we have shown above that the probability of any c_i ringing exactly once is less than or equal to $2M\tau$. Since $T_L < 1$, the expected energy gain associated with E_{1b} is less than $\mathbb{P}[E_{1b}]$, and $2M^2\tau^2 \le 2e^{-2}Ch(1)\tau$.

Event E_k, $k>1$: The times at which exchanges with baths take place are given by a rather complicated process: Consider, for example, c_0. At any one moment, the time to the next ring is given by an exponential distribution, but the parameters of this distribution vary with x_1, the value of which is in turn dependent on the history of interactions in the entire chain. When c_0 rings, the sequence of energies emitted by the bath is *i.i.d.*, however, and we will take advantage of that.

To put the random variables representing energies emitted by the left and right baths on the same probability space, we introduce

$$\varphi_\theta : (0,1) \to (0,\infty) \quad \text{defined by } \int_0^{\varphi_\theta(x)} \frac{1}{\theta} e^{-y/\theta} \mathrm{d}y = x,$$

so that φ_θ carries the uniform distribution on the unit interval $(0,1)$ to the exponential distribution with mean θ.

We assume throughout that $\mathbf{x}(0)$ is the initial condition and $[0,\tau)$ is the time interval of interest. First fix $\mathbf{u} = (u_1, \ldots, u_k) \in (0,1)^k$, and define $\Omega_{\mathbf{u}}$ to be the set of sample paths ω with the property that at the ith energy exchange with a bath, $1 \le i \le k$, the energy drawn from the bath (to be mixed with x_1 or x_N) is $\varphi_{\sigma(\omega,i)}(u_i)$, where $\sigma(\omega, i) = T_L$ (resp. T_R) if this exchange is with the left (resp. right) bath. Thus for this sample path, the total amount of energy to enter the chain in the first k exchanges with baths is bounded above by

$$\sum_{i=1}^k \varphi_{\sigma(\omega,i)}(u_i) \le \sum_{i=1}^k \varphi_1(u_i) := S_{k,\mathbf{u}}.$$

In the last inequality, we have used the fact that for fixed $x \in (0,1)$, $\varphi_\theta(x)$ increases with θ. Let $\mathbb{P}_{\mathbf{u}}$ denote the conditional probability of $\mathbb{P}$ on $\Omega_{\mathbf{u}}$, and $\mathbb{E}_{\mathbf{u}}$ the corresponding expectation.

We claim that

$$\mathbb{P}_{\mathbf{u}}(E_k) = \int \cdots \int_{\{0<s_1<s_2<\cdots<s_k<\tau\}} p_1 q_1 p_2 q_2 \cdots p_k q_k p_{k+1} \mathrm{d}s_1 \cdots \mathrm{d}s_k,$$

where p_i and $q_i ds_i$ are the probabilities of events P_i and Q_i:

$$P_i = \big\{\text{neither } c_0 \text{ nor } c_N \text{ rings on } (s_{i-1}, s_i) \mid P_j, Q_j, j < i\big\};$$
$$Q_i = \big\{\text{one of } c_0 \text{ or } c_N \text{ rings on } (s_i, s_i + ds_i) \mid P_j, j \le i, \text{ and } Q_j, j < i\big\}.$$

The definitions above are to be read with $s_0 = 0$ and $s_{k+1} = 1$. Using the generous bounds $p_i \le 1$ and

$$q_i \le f\big(x_1(s_i), T_L\big) + f\big(x_N(s_i), T_R\big) \le 2\big(M + Ch(1) S_{k,\mathbf{u}}\big),$$

we obtain

$$\mathbb{P}_{\mathbf{u}}(E_k) \le \frac{\tau^k}{k!} \big[2\big(M + Ch(1) S_{k,\mathbf{u}}\big)\big]^k.$$

With respect to $\mathbb{E}_{\mathbf{u}}$, then, the expected energy gain associated with E_k is no greater than

$$\begin{aligned} \mathbb{E}_{\mathbf{u}}\left[\left(\sum_{i=1}^k \varphi_{\sigma(\cdot,i)}(u_i)\right) \mathbf{1}_{E_k}\right] &\le S_{k,\mathbf{u}} \cdot \mathbb{P}_{\mathbf{u}}[E_k] \\ &\le S_{k,\mathbf{u}} \frac{2^k \tau^k}{k!} \big(M + Ch(1) S_{k,\mathbf{u}}\big)^k. \end{aligned}$$

Integrating over $\mathbf{u}$, we obtain the upper bound

$$\mathbb{E}\left[S_k \frac{2^k \tau^k}{k!} \big(M + Ch(1) S_k\big)^k\right] \tag{5.1}$$

where $S_k = \sum_{i=1}^k Y_i$ is the sum of k independent, mean-1, exponentially distributed, random variables (called an Erlang distribution).

The probability density function of S_k is $x^{k-1}e^{-x}/(k-1)!$, and its moment generating function is $(1-t)^{-k}$. Differentiating the latter m times, we obtain

$$\mathbb{E}\big[S_k^m\big] = k(k+1)\cdots(k+m-1) < \prod_{j=0}^{m-1}(k+j).$$

Thus the expectation in (5.1) is equal to

$$\begin{aligned}\frac{2^k\tau^k}{k!}\sum_{i=0}^{k}\binom{k}{i}M^{k-i}\big(Ch(1)\big)^i\mathbb{E}\big[S_k^{i+1}\big] &< \frac{2^k\tau^k}{k!}\sum_{i=0}^{k}\binom{k}{i}M^{k-i}\big(Ch(1)\big)^i(2k)^{i+1}\\ &= \frac{2k(2^k\tau^k)}{k!}\big(M+2Ch(1)k\big)^k.\end{aligned}$$

This completes the estimate for the contribution from one E_k at a time, for $k \geq 2$. (The argument is valid for $k = 1$, but the bound is too weak to be useful.) To finish, we consider the expected energy gain from $\bigcup_{k\geq 2} E_k$. By Stirling's formula,

$$n! \geq \sqrt{2\pi n}\frac{n^n}{e^n}$$

for all n (see [32] for the version used). Hence

$$\begin{aligned}\sum_{k=2}^{\infty}\frac{2k(2^k\tau^k)}{k!}\big(M+2Ch(1)k\big)^k &\leq \sum_{k=2}^{\infty}\frac{2ke^k}{\sqrt{2\pi k}\cdot k^k}\cdot(2kM\tau)^k\left(\frac{1}{k}+\frac{2Ch(1)}{M}\right)^k\\ &< \sum_{k=2}^{\infty}\frac{2ke^k}{\sqrt{2\pi k}\cdot k^k}\cdot(2kM\tau)^k\cdot e \quad \text{since } 2Ch(1)\leq M\\ &\leq \sum_{k=2}^{\infty}\frac{2ke}{\sqrt{4\pi}}(2M\tau e)^k.\end{aligned}$$

Letting $a := 2M\tau e$ and noticing that $a \leq e^{-1} < \frac{1}{2}$, we see that the sum above is

$$= \frac{e}{\sqrt{\pi}}\sum_{k=2}^{\infty}ka^k = \frac{e}{\sqrt{\pi}}\frac{a^2(2-a)}{(1-a)^2} < \frac{2e}{\sqrt{\pi}}a^2 = \frac{8e}{\sqrt{\pi}}Ch(1)\tau.$$

The finiteness of this sum implies in particular that $\mathbb{P}[E_\infty] = 0$.

Summing the contribution from all cases, we obtain that $G_{[0,\tau)}(\mathbf{x}(0))$ is bounded above by $B\tau$ for a constant B independent of $\mathbf{x}(0)$. □

5.2 Proof of Proposition 5.1

Starting from $\mathbf{x}(0)$, we introduce two sequences of random variables $\tau_1, \tau_2, \dots$ and $X_0, X_1, \dots$ as follows. To begin with, we let $X_0 = \max\{2, V_0(\mathbf{x}(0))\}Ch(1)$ and $\tau_1 =$

$Ch(1)e^{-2}X_0^{-2}$ (cf. Lemma 5.2). Let X_1 be the energy gain on the time interval $[0, \tau_1)$ in the sense of the last section, i.e.,

$$X_1 = \sum_{t_\alpha} \max\{x_1(t_\alpha^+) - x_1(t_\alpha), 0\}\mathbf{1}_{\{0<t_\alpha<\tau_1\}} + \sum_{s_\alpha} \max\{x_N(s_\alpha^+) - x_N(s_\alpha), 0\}\mathbf{1}_{\{0<s_\alpha<\tau_1\}},$$

where t_α (resp. s_α) are the times at which c_0 (resp. c_N) rings.

By Lemma 5.2, $\mathbb{E}_{\mathbf{x}(0)}[X_1] \le B\tau_1$. Since $\mathbf{x}(0)$ is fixed throughout, we will drop the subscript in $\mathbb{E}$ from here on.

Having defined τ_k and X_k for all $k < n$, we define recursively

$$\tau_n = \frac{Ch(1)e^{-2}}{(X_0 + Ch(1)\sum_{1\le k<n} X_k)^2}, \tag{5.2}$$

and let X_n be the energy gained on the "next" τ_n units of time, i.e. on the time interval $[t_{n-1}, t_n)$ where $t_n = \sum_{i\le n} \tau_i$.

Observe that Lemma 5.2 can be applied to give

$$\mathbb{E}[X_n \mid \tau_n = s] \le Bs,$$

for if $\tau_n = s$, then $V_0(t_{n-1})$, the total energy of the system at time t_{n-1}, satisfies

$$\max\{2, V_0(t_{n-1})\}Ch(1) \le X_0 + Ch(1) \sum_{1\le k<n} X_k.$$

(The inequality is likely strict, for X_k counts only (positive) net gain per interaction ignoring net losses.) Repeated applications of Lemma 5.2 then gives

$$\mathbb{E}\left[\sum_{k=1}^{n} X_k \,\middle|\, \sum_{k=1}^{n} \tau_k = s\right] \le Bs,$$

and if $\mathbb{P}[\sum_{k=1}^{\infty} \tau_k \le 1] = 0$, then we have proved

$$G_{[0,1)}(\mathbf{x}(0)) \le B.$$

It remains, therefore, to show $\mathbb{P}[\sum_{k=1}^{\infty} \tau_k \le 1] = 0$. Suppose not. Then for each sample path with $\sum_{k=1}^{\infty} \tau_k \le 1$, there is an infinite sequence $k_1 < k_2 < \cdots$ with the property that $\tau_{k_j} < 1/k_j$; the absence of such a sequence would imply $\tau_k \ge 1/k$ for all but finitely many k, leading to $\sum_{k=1}^{\infty} \tau_k = \infty$. But $\tau_{k_j} < 1/k_j$ implies, by (5.2), that

$$\sum_{i=1}^{k_j-1} X_i > \frac{\sqrt{k_j Ch(1)}e - X_0}{Ch(1)},$$

which tends to infinity as $k_j \to \infty$, contradicting

$$\mathbb{E}\left[\sum_{k=1}^{\infty} X_k \,\middle|\, \sum_{k=1}^{\infty} \tau_k \le 1\right] < B.$$

This completes the proof of Proposition 5.1.

6 Relevant Properties of the Time-1 Chain

In Sects. 6.1 and 6.2, we show that Φ_n, the time-1 sample chain of Φ_t, satisfies conditions Assumptions (A1) and (A2) in Sect. 4.1. These are the main ingredients in the proofs of Theorems 1 and 3. Section 6.3 discusses how Φ_1 transforms the Lebesgue measure class, an issue at the heart of Theorem 2.

6.1 Total Energy as Lyapunov Function

In analogy with the definitions related to expected energy gain at the beginning of Sect. 5, we define expected energy loss to be $L := L_L + L_R$ where

$$L_{L,[0,\tau)}\big(\mathbf{x}(0)\big) := \mathbb{E}_{\mathbf{x}(0)}\left[\sum_{t_\alpha} \max\big\{x_1(t_\alpha) - x_1\big(t_\alpha^+\big), 0\big\}\mathbf{1}_{t_\alpha<\tau}\right] \tag{6.1}$$

where $t_\alpha > 0$ are the times when c_0 rings, and $L_{R,[0,\tau)}(\mathbf{x}(0))$ is defined analogously.

Lemma 6.1 *There exist $\gamma_0 > 0$ and $K_* < \infty$ such that for every $\mathbf{x}(0)$ with $V_0(\mathbf{x}(0)) > K_*$,*

$$L_{[0,1)}\big(\mathbf{x}(0)\big) \geq \gamma_0 V_0\big(\mathbf{x}(0)\big).$$

Proof We first lay out a procedure that we claim will lead to the systematic dumping of a fraction of $V_0(\mathbf{x}(0))$ into one of the baths – postponing the choice of constants and computation of probabilities till later.

Let $M = x_{n_0}(0) = \max\{x_i(0), 1 \leq i \leq N\}$, i.e., initially, site n_0 has the highest energy among all sites. Of interest to us is the event

$$\mathbf{S} = \bigcap_{i=1}^{n_0} S_i$$

where S_i defines what happens on the time interval $[(i-1)\delta, i\delta)$ and the size of δ is to be specified:

$$S_i = \Big\{\text{The following hold on the time interval } [(i-1)\delta, i\delta):$$

(a) c_{n_0-i} rings exactly once;

(b) all other c_j are silent; and

(c) $x_{n_0-i}(i\delta) \geq \dfrac{1}{2} x_{n_0-i+1}\big((i-1)\delta\big)\Big\}$.

That is to say, S_1 defines what takes place on $[0, \delta)$. In this event, c_{n_0-1} is the only clock that rings; it rings exactly once, say at time s_1, resulting in $x_{n_0-1}(s_1^+) \geq \frac{1}{2}x_{n_0}(s_1) = \frac{1}{2}x_{n_0}(0) \geq \frac{1}{2}M$. The next ring, which takes place at time $s_2 \in [\delta, 2\delta)$, comes from c_{n_0-2}, and results in $x_{n_0-2}(s_2^+) \geq \frac{1}{2}x_{n_0-1}(s_2) \geq \frac{1}{4}M$. Inductively, one sees that $x_1(s_{n_0-1}^+) \geq 2^{-(n_0-1)}M$. Definitions associated with S_{n_0} have to be interpreted a little differently: $x_0(s_{n_0})$ is the energy emitted by the left bath in the exchange at time s_{n_0}, and $x_0(s_{n_0}^+)$ is the portion of $x_0(s_{n_0}) + x_1(s_{n_0})$ that goes to the bath. Event S_{n_0} again stipulates that $x_0(s_{n_0}^+) \geq \frac{1}{2}x_1(s_{n_0}) \geq$

$2^{-n_0}M$. The amount of energy dumped, i.e., lost by the chain in the sense of (6.1), is $\max\{x_0(s_{n_0}^+) - x_0(s_{n_0}), 0\}$, and this is guaranteed to happen with probability $\mathbb{P}[\mathbf{S}]$.

Observe that we may assume $n_0 \le N/2$ in the scheme above, for if $n_0 > N/2$, we can pass the energy to the right bath. It is clear from the discussion above that the sample paths in $\mathbf{S}$ lead to the dumping of a fraction of the highest energy. To obtain the desired bound on expected energy loss, we need to show that $\mathbb{P}[\mathbf{S}]$ is bounded below by a quantity independent of $V_0(\mathbf{x}(0))$. This is what the rest of the proof is about. As the computation is slightly different depending on the relation between interaction rates and system size, we separate into the following two cases:

Case 1. $N < \sup_x h(x)$. We let M_0 be such that $h(M_0) \ge \frac{N}{2}$, and consider only those initial conditions with $V_0(\mathbf{x}(0)) \ge NM_0$, so that M, the highest energy in any one site initially, is $\ge M_0$. We choose our time steps to be $\delta = h(M)^{-1}$, and notice that the process is completed within one unit of time, since $n_0\delta \le \frac{N}{2}h(M)^{-1} \le \frac{N}{2}h(M_0)^{-1} \le 1$ by the nondecreasing property of h.

We write

$$\mathbb{P}[S_i \mid S_1, \dots, S_{i-1}] = P_1(i)P_2(i)P_3(i)$$

where P_1 and P_2 correspond respectively to the probabilities of (a) and (b) in the definition of S_i given S_j for all $j < i$, and P_3 is the probability of (c) given (a), (b) and S_j, $j < i$. Observe that given $S_1, \dots, S_{i-1}$, we have, at time $(i-1)\delta$,

(i) $x_{n_0-i} = x_{n_0-i}(0) \le M$, and
(ii) $2^{-i+1}M \le x_{n_0-i+1} \le iM$.

(i) is clear, since neither of the clocks adjacent to it has rung so far. The lower bound in (ii) is a consequence of S_j, $j < i$, and the upper bound is attained if $x_j(0) = M$ for every $j \in [n_0 - i + 1, n_0]$, and in every one of the first $i - 1$ interactions, all of the energy goes to the site on the left.

To estimate $P_1(i)$, we let x_j denote the energy at site j before the clock rings and $\bar{x}_j$ after. Then, using the nondecreasing property and the sublinear growth rate of h, we obtain

$$\begin{aligned}
P_1(i) &= \int_0^\delta f(x_{n_0-i}, x_{n_0-i+1})e^{-f(x_{n_0-i},x_{n_0-i+1})s}e^{-f(\bar{x}_{n_0-i},\bar{x}_{n_0-i+1})(\delta-s)}\mathrm{d}s \\
&\ge \int_0^\delta h\left(\frac{M}{2^{i-1}}\right)e^{-Ch(iM)s}e^{-Ch((i+1)M)(\delta-s)}\mathrm{d}s \\
&= h\left(\frac{M}{2^{i-1}}\right) \cdot e^{-Ch((i+1)M)\delta} \cdot \delta \\
&\ge 2^{-(i-1)}h(M) \cdot e^{-C(i+1)h(M)\delta} \cdot \delta \\
&= 2^{-(i-1)}e^{-C(i+1)} \quad \text{since } \delta h(M) = 1.
\end{aligned}$$

Likewise,

$$P_2(i) = \prod_{j \ne n_0 - i} e^{f(x_j,x_{j+1})\delta} \ge \left(e^{-Ch(iM)\delta}\right)^N \ge e^{-CiN},$$

using again $\delta h(M) = 1$, and it is obvious that $P_3(i) \ge \frac{1}{2}$. These estimates together show that $\mathbb{P}[\mathbf{S}] = \prod_{i=1}^{n_0} P_1(i)P_2(i)P_3(i)$ is bounded below by a quantity independent of $V_0(\mathbf{x}(0))$.

Finally, given $\mathbf{S}$, the expected energy loss at time s_{n_0} is greater than $\frac{1}{2}x_0(s_{n_0}^+) \geq 2^{-n_0-1}M$ if $x_0(s_{n_0}^+) \geq 2$ (since $\mathbb{E}[x_0(s_{n_0})] = T_L < 1$). Thus if we take $K_* = N \cdot \max\{M_0, 2^{\frac{N}{2}+1}\}$, then $M \geq \max\{M_0, 2^{n_0+1}\}$, $x_0(s_{n_0}^+) \geq 2$, and the assertion in Lemma 6.1 holds with $\gamma_0 = \mathbb{P}[\mathbf{S}]2^{-(\frac{N}{2}+1)}N^{-1}$.

Case 2. $N \geq \sup_x h(x)$. When it is not possible to satisfy $h(M)\delta = 1$ and $N\delta \leq 1$ at the same time, we proceed a little differently: We let $h_0 := \sup_x h(x)$, fix M_0 with $h(M_0) \geq \frac{1}{2}h_0$, and choose $\delta = \frac{1}{N}$. Then for $M \geq M_0$, we have $\frac{1}{N}\frac{h_0}{2} \leq \delta h(M) \leq \frac{1}{N}h_0$, which is sufficient for our purposes. □

We now state for the record

Proposition 6.2 *The time-*1 *chain Φ_n satisfies Assumption (A1) with $W_0 = V_0 = \sum_{i=1}^N x_i$, i.e., there exist $K_0 < \infty$ and $\gamma_0 > 0$ such that for all $\mathbf{x} \in \mathbb{R}_+^N$, we have*

$$(PV_0)(\mathbf{x}) - V_0(\mathbf{x}) \leq -\gamma_0 V_0(\mathbf{x}) + K_0.$$

Proof Since

$$PV_0(\mathbf{x}) - V_0(\mathbf{x}) = G_{[0,1)}(\mathbf{x}) - L_{[0,1)}(\mathbf{x})$$

where $G_{[0,1)}$ and $L_{[0,1)}$ are the expected energy gain and loss in one unit of time for the continuous-time chain Φ_t, it follows from Proposition 5.1 and Lemma 6.1 that

$$(PV_0)(\mathbf{x}) - V_0(\mathbf{x}) \leq -\gamma_0 V_0(\mathbf{x}) + B$$

whenever $V_0(\mathbf{x}) \geq K_*$. The desired inequality then holds for all $\mathbf{x} \in \mathbb{R}_+^N$ with $K_0 = B + \gamma_0 K_*$. □

6.2 Doeblin-Type Condition on Sets with Bounded Total Energy

In what follows we will use the notation $D_a := \{V_0 \leq a\}$, and $B_a := \{\mathbf{x} : x_i < a\ \forall i\}$. Recall that Lebesgue measure on $\mathbb{R}^N$ is denoted by λ, and for $B \subset \mathbb{R}_+^N$, $\lambda|_B$ is the restriction of λ to B.

Lemma 6.3 *Given any $M, a > 0$, there exists $\epsilon = \epsilon(M, a) > 0$ such that*

$$P(\mathbf{x}, \cdot) \geq \epsilon(\lambda|_{B_a}) \quad \textit{for all } \mathbf{x} \in D_M. \tag{6.2}$$

Proof For $\mathbf{z} = (z_1, \dots, z_N) \in \mathbb{R}_+^N$ and $\sigma > 0$, let $B(\mathbf{z}, \sigma) := \{z_i \leq x_i < z_i + \sigma,\ \text{all } i\}$. It suffices to produce, for given $M, a > 0$, an ϵ for which

$$P\big(\mathbf{x}, B(\mathbf{z}, \sigma)\big) \geq \epsilon\lambda\big(B(\mathbf{z}, \sigma)\big)$$

holds for all $\mathbf{x} \in D_M$ and all $\mathbf{z}, \sigma$ such that $B(\mathbf{z}, \sigma) \subset B_a$. Let $M, a, \mathbf{x}, \mathbf{z}$ and σ be fixed.

As before, we first identify the event of interest: Let $\delta = \frac{1}{N+1}$. We consider $\mathbf{S} = S_0 \cdots S_N$ where S_i, which specifies what happens on the time interval $[i\delta, (i+1)\delta)$, is defined as follows:

$$S_i = A_i \cap \big\{\text{on } [i\delta, (i+1)\delta),\ c_i \text{ rings exactly once, all other clocks are silent}\big\};$$

$$A_0 = \big\{\text{energy emitted by left bath } \in (Na+1, Na+2)\big\},$$

$$A_i = \{x_i((i+1)\delta) \in [z_i, z_i + \sigma)\} \quad \text{for } i = 1, 2, \ldots, N.$$

The idea is that a sufficiently large amount of energy is injected into the chain initially so that for $i = 1, \ldots, N$, $x_i(i\delta) \geq 1 + (N - i + 1)a$, so it is always possible for site i to acquire an amount of energy between z_i and $z_i + \sigma$ and to pass the rest to site $i + 1$, or to the right bath in the case $i = N$.

We seek an estimate from below in the form of $\epsilon\sigma^N$ for $\mathbb{P}[\mathbf{S}]$. The situation is simpler than in Lemma 6.1 as ϵ is allowed to depend on M and a, as well as N, h, T_L and T_R. Here are the considerations:

(i) At time $i\delta$, we need a lower bound on $\eta_i =$ the rate of c_i, to bound from below the probability of c_i ringing within time δ. Here we have $\eta_0 \geq h(T_L)$, and for $i = 1, \ldots, N$, $\eta_i = h(\max\{x_i(i\delta), x_{i+1}(i\delta)\}) \geq h(1 + a)$ conditioned on S_j, $j < i$.
(ii) There is a uniform upper bound for η_i as $V_0 \leq M + Na + 2$ throughout.
(iii) Finally, let $p \in (0, 1)$ be the fraction in the next mixing that puts $x_i \in [z_i, z_i + \sigma)$. Given S_j, $j < i$, we will need $p \geq c\sigma$ for some $c > 0$. For $i < N$, this is true as $1 + a \leq x_i(i\delta) + x_{i+1}(i\delta) \leq M + Na + 2$. The case of $i = N$ is similar.

Further details are left to the reader. □

It follows immediately from Lemma 6.3 that Assumption (A2) is satisfied by the time-1 chain Φ_n. We have shown, in fact, that D can be taken to be $\{V_0 \leq M\}$ for any M, and ν can be taken to be $\frac{1}{\lambda(B_a)}\lambda|_{B_a}$ for any $a > 0$. We record the following immediate corollary, which connects the previous discussion to the next:

Corollary 6.4 *For every $\mathbf{x} \in \mathbb{R}_+^N$, $P(\mathbf{x}, \cdot)$ has an absolutely continuous component with a strictly positive density.*

We clarify this statement: By the Lebesgue Decomposition Theorem, any finite Borel measure μ on $\mathbb{R}_+^N$ can be decomposed into $\mu = \mu_{\text{abs}} + \mu_\perp$ where μ_{abs} is absolutely continuous with respect to λ (written $\mu_{\text{abs}} \ll \lambda$) and $\mu_\perp$ is singular with respect to λ (written $\mu_\perp \perp \lambda$). Letting $\mu = P(\mathbf{x}, \cdot)$, we observe that both μ_{abs} and $\mu_\perp$ are nontrivial for every $\mathbf{x}$: Corollary 6.4 tells us that not only is $\mu_{\text{abs}}(\mathbb{R}_+^N) > 0$, the Radon–Nikodym derivative of μ_{abs} with respect to λ (written $\frac{d\mu_{\text{abs}}}{d\lambda}$) is > 0 λ-a.e. To see that $\mu_\perp$ is nontrivial, notice that there is always a positive measure set of sample paths that correspond, for example, to c_{i-1} and c_i not ringing before time 1 for some i. Following these sample paths, the value of x_i is not disturbed.

6.3 Action of Markov Operator on Lebesgue Measure Class

Recall that associated with the time-1 chain, we have the operator P acting on measures by

$$(\mu P)(\cdot) = \int P(\mathbf{x}, \cdot)\mu(d\mathbf{x}).$$

Lemma 6.5 *For any finite Borel measure μ on $\mathbb{R}_+^N$, if $\mu \ll \lambda$, then $\mu P \ll \lambda$.*

Proof Let

$$\mathcal{L}(N) = \{\ell = (i_1, i_2, \ldots, i_n) \mid 0 \leq i_1, i_2, \ldots, i_n \leq N, n \in \mathbb{Z}^+\}.$$

For $\ell = (i_1, i_2, \dots, i_n)$, we define the event $E(\ell)$ to be that in which clocks $c_{i_1}, c_{i_2}, \dots, c_{i_n}$ ring in the order specified and these are only energy exchanges before $t = 1$. We also let $E(\emptyset)$ and $E(\infty)$ denote respectively the event of zero or infinitely many rings on the time interval $[0, 1)$. For $E = E(\ell)$, $E(\emptyset)$ or $E(\infty)$, we define the operator P_E, which acts on the space of Borel measures on $\mathbb{R}^N_+$, by

$$(\mu P_E)(A) = \int P_E(\mathbf{x}, A)\mu(\mathrm{d}\mathbf{x}) \quad \text{where } P_E(\mathbf{x}, A) = \mathbb{P}\big[(\Phi_1 \in A \mid \Phi_0 = \mathbf{x}) \mid E\big].$$

Then

$$\mu P(\cdot) = \int_{\mathbb{R}^N_+} \Bigg[\Bigg(\sum_{\ell \in \mathcal{L}(N)} \mathbb{P}_{\mathbf{x}}\big[E(\ell)\big] P_{E(\ell)}(\mathbf{x}, \cdot)\Bigg) + \mathbb{P}_{\mathbf{x}}\big[E(\emptyset)\big] P_{E(\emptyset)}(\mathbf{x}, \cdot)$$

$$+ \mathbb{P}_{\mathbf{x}}\big[E(\infty)\big] P_{E(\infty)}(\mathbf{x}, \cdot)\Bigg]\mu(\mathrm{d}\mathbf{x})$$

We will check at the end of the proof that $\mathbb{P}_{\mathbf{x}}[E(\infty)] = 0$ for every $\mathbf{x}$. Focus first on the following two crucial observations:

(1) Given $\mu \ll \lambda$ and $\ell \in \mathcal{L}(N)$, we let μ_ℓ be the measure on $\mathbb{R}^N_+$ with the property that

$$\frac{d\mu_\ell}{d\lambda}(\mathbf{x}) = \mathbb{P}_{\mathbf{x}}\big[E(\ell)\big]\frac{d\mu}{d\lambda}(\mathbf{x}).$$

Notice that $0 < \mu_\ell(\mathbb{R}^N_+) < 1$. Similarly define $\mu_\emptyset$. Then

$$\mu P = \mu_\emptyset + \sum_{\ell \in \mathcal{L}(N)} \mu_\ell P_{E(\ell)}, \tag{6.3}$$

and it suffices to show the absolute continuity of each of the countably many measures on the right side of (6.3).

(2) For every $\ell = (i_1, \dots, i_n)$, the operator $P_{E(\ell)}$ can be decomposed into

$$P_{E(\ell)} = P_{E(i_n)} \cdots P_{E(i_2)} P_{E(i_1)},$$

where $E(k)$ is the event that on the time interval $[0, 1)$, c_k rings exactly once, and all other clocks are silent.

Observations (1) and (2) together reduce the problem to the following: Given $\mu \ll \lambda$ and $k \in \{0, 1, \dots, N\}$, it suffices to show that $\mu P_{E(k)} \ll \lambda$. Let $\xi = \frac{d\mu}{d\lambda}$. We claim, and leave the verification (which is straightforward) to the reader, that $\mu P_{E(k)}$ is the measure whose density η_k with respect to λ is given by

$$\eta_k(\mathbf{x}) = \int_0^1 \xi\big(x_1, \dots, x_{k-1}, p(x_k + x_{k+1}), (1-p)(x_k + x_{k+1}), x_{k+2}, \dots, x_N\big)\mathrm{d}p$$

for $0 < k < N$ and $\mathbf{x} = (x_1, \dots, x_N)$, while for $k = 0$ and N, we have

$$\eta_0(\mathbf{x}) = \int_0^\infty \int_{(x_1 - \hat{x}_1)^+}^\infty \xi(\hat{x}_1, x_2, \dots, x_N)\frac{1}{\hat{x}_1 + y}\beta_L e^{-\beta_L y}\mathrm{d}y\mathrm{d}\hat{x}_1,$$

$$\eta_N(\mathbf{x}) = \int_0^\infty \int_{(x_N - \hat{x}_N)^+}^\infty \xi(x_1, x_2, \dots, \hat{x}_N)\frac{1}{\hat{x}_N + y}\beta_R e^{-\beta_R y}\mathrm{d}y\mathrm{d}\hat{x}_N.$$

We finish with a proof of $\mathbb{P}_{\mathbf{x}}[E(\infty)] = 0$. Let $\mathbf{x} \in \mathbb{R}_+^N$ be fixed. Given $\varepsilon > 0$, we know from Proposition 5.1 that for $\bar{M}$ large enough, $\mathbb{P}_{\mathbf{x}}[F_{\bar{M}}] > 1 - \varepsilon$ where $F_{\bar{M}} = \{V_0(\mathbf{x}(t)) \le \bar{M}, t \in [0,1)\}$. With total energy $\le \bar{M}$, each c_i is, at any one point in time, no faster than a clock with rate $\bar{M}$. An estimate similar to that of $\mathbb{P}_{\mathbf{u}}(E_k)$ in the proof of Lemma 5.2 gives

$$\mathbb{P}_{\mathbf{x}}[\text{exactly } k \text{ rings} \mid F_{\bar{M}}] \le \frac{\bar{M}^k}{k!}.$$

From this one deduces that

$$\mathbb{P}_{\mathbf{x}}[\ge n \text{ rings} \mid F_{\bar{M}}] \le \sum_{k \ge n} \frac{((N+1)\bar{M})^k}{k!},$$

which tends to 0 as $n \to \infty$. □

7 Proof of Theorems

7.1 Summary of Results for Time-Δ Sample Chains

Recall that for any $\Delta > 0$, the time-Δ sample chain of Φ_t is denoted by Φ_n^Δ. We summarize here a few results for the discrete-time chain Φ_n^Δ the main computations for which have been carried out in Sects. 5 and 6. Let $V(\mathbf{x}) = V_0(\mathbf{x}) + 1$.

Proposition 7.1

(a) *For every* $\Delta > 0$, $\exists$ *a unique* $\pi_\Delta \in L_V(\mathbb{R}_+^N)$ *left invariant by* Φ_n^Δ.
(b) *There exist constants* $C_0 > 0$ *and* $\rho_0 \in (0,1)$ *such that for all* $\Delta \in [1,2]$,

$$\left\| \mu_1 P^{n\Delta} - \mu_2 P^{n\Delta} \right\|_V \le C_0 \rho_0^n \|\mu_1 - \mu_2\|_V$$

holds for all $\mu_1, \mu_2 \in L_V(\mathbb{R}_+^N)$.

Proof For $\Delta = 1$, Assumption (A1) is given by Proposition 6.2, and Assumption (A2) is proved in Sect. 6.2. These results are extended *mutatis mutandis* to any $\Delta > 0$. Thus part (a) in the proposition follows immediately from Theorem 4.1, as does part (b) if constants are permitted to depend on Δ. That two numbers C_0 and ρ_0 can be chosen to work for all $\Delta \in [1,2]$ is what remains to be checked. We leave it to the reader to verify that uniform constants can be chosen for the range of Δ specified in the proofs of Proposition 6.2 and Lemma 6.3. (The choice of $[1,2]$ is arbitrary but upper and lower bounds are clearly needed: For example, expected energy gain on the time interval $[0,\Delta)$ is shown to be less than $B\Delta$, and the constant ϵ in Lemma 6.3 goes to zero as $\Delta \to 0$.) Explicit relationships between the constants in Assumption (A1) and (A2) and the numbers C_0 and ρ_0 are given in item (2) in the Remark following Theorem 4.1. □

Proposition 7.2 *The following hold for any* $\Delta > 0$:

(a) *Every* Φ_n^Δ*-invariant Borel probability measure* μ *is* $\ll \lambda$ *with* $\frac{d\mu}{d\lambda} > 0$ λ*-a.e.*
(b) *It follows from* (a) *that* Φ_n^Δ *has at most one invariant probability measure.*

Proof The proofs are identical for all $\Delta > 0$. For notational simplicity, we assume $\Delta = 1$.

(a) Let μ be an invariant probability measure for Φ_n, and let $\mu = \mu_{\text{abs}} + \mu_\perp$ be the decomposition in the last paragraph of Sect. 6.2. Then $\mu P = \mu_{\text{abs}} P + \mu_\perp P$. By Lemma 6.5, we have $\mu_{\text{abs}} P \ll \lambda$, while Corollary 6.4 tells us that if $\mu_\perp(\mathbb{R}^N_+) \neq 0$, then $\mu_\perp P$ will have a nonzero absolutely continuous component. Since it is impossible for $((\mu P)_\perp)(\mathbb{R}^N_+)$ to be strictly smaller than $(\mu_\perp)(\mathbb{R}^N_+)$, we conclude that $\mu_\perp(\mathbb{R}^N_+) = 0$, i.e., $\mu \ll \lambda$. That $\frac{d\mu}{d\lambda} > 0$ λ-a.e also follows from Corollary 6.4.

(b) is an immediate consequence of (a) as there cannot be two distinct ergodic measures both with positive densities λ-a.e. □

7.2 From Discrete to Continuous Time

We need to show that all the π_Δ given by Proposition 7.1(a) are identical, and to do that requires the following "continuity at 0" property for the family of time-δ sample chains.

Lemma 7.3 *Given $\Delta > 0$ and $\epsilon > 0$, there exists $\delta_0 = \delta_0(\Delta, \epsilon) > 0$ such that for all $\delta \in (0, \delta_0)$,*

$$\left|\left(\pi_\Delta P^\delta\right)(A) - \pi_\Delta(A)\right| < \epsilon$$

for every Borel set $A \subset \mathbb{R}^N_+$.

Proof For fixed Δ, ϵ, there exists a compact set $U = \{\mathbf{x} \mid x_i \leq R\}$ such that $\pi_\Delta(U) > 1 - \frac{\epsilon}{4}$. Since the rate function $f(x, y)$ is bounded on U, there is a real number $\delta_0 > 0$ such that $P^\delta(\mathbf{x}, \mathbf{x}) > 1 - \frac{\epsilon}{4}$ (e.g., no energy exchange takes place on the time interval $(0, \delta)$) for every $\mathbf{x} \in U$ and $0 < \delta < \delta_0$. Then we have

$$\begin{aligned}
\left(\pi_\Delta P^\delta\right)(A) &= \int P^\delta(\mathbf{x}, A)\pi_\Delta(d\mathbf{x}) \\
&= \int_{U\cap A} P^\delta(\mathbf{x}, A)\pi_\Delta(d\mathbf{x}) + \int_{U-A} P^\delta(\mathbf{x}, A)\pi_\Delta(d\mathbf{x}) + \int_{U^c} P^\delta(\mathbf{x}, A)\pi_\Delta(d\mathbf{x}) \\
&= \pi_\Delta(U \cap A) - a_1 + a_2 + a_3
\end{aligned}$$

where

$$\begin{aligned}
a_1 &= \int_{U\cap A} \left(1 - P^\delta(\mathbf{x}, A)\right)\pi_\Delta(d\mathbf{x}) \leq \frac{\epsilon}{4}\pi_\Delta(U \cap A) \leq \frac{\epsilon}{4}, \\
a_2 &= \int_{U-A} P^\delta(\mathbf{x}, A)\pi_\Delta(d\mathbf{x}) \leq \frac{\epsilon}{4}\pi_\Delta(U - A) \leq \frac{\epsilon}{4}, \\
a_3 &= \int_{U^c} P^\delta(\mathbf{x}, A)\pi_\Delta(d\mathbf{x}) \leq \pi_\Delta\left(U^c\right) \leq \frac{\epsilon}{4}.
\end{aligned}$$

Further, $\pi_\Delta(A) - \pi_\Delta(U \cap A) \leq \pi_\Delta(U^c) < \epsilon/4$, so $|\pi_\Delta(A) - \pi_\Delta(U \cap A)| \leq \epsilon/4$. Hence

$$\left|\pi_\Delta(A) - \left(\pi_\Delta P^\delta\right)(A)\right| < \epsilon.$$

□

Proof of Theorem 1 To show the existence of $\pi \in L_V(\mathbb{R}^N_+)$ with $\pi P^t = \pi$ for all $t > 0$, we need to show $\pi_\Delta = \pi_\Gamma$ for all $\Delta, \Gamma > 0$. That there is at most one such measure follows from the corresponding result for Φ^Δ_n.

Let Δ and Γ be fixed. It suffices to show for an arbitrarily small (but fixed) $\epsilon > 0$ that

$$\left|\left(\pi_\Delta P^\Gamma\right)(A) - \pi_\Delta(A)\right| < \epsilon \tag{7.1}$$

holds for every Borel set $A \subset \mathbb{R}^N_+$.

Notice first that for every $j \in \mathbb{Z}^+$, since π_Δ is also invariant for the time-$j\Delta$ sample chain $\Phi_n^{j\Delta}$, we have $\pi_{j\Delta} = \pi_\Delta$. Similarly, for any $k \in \mathbb{Z}^+$, $\pi_{j\Delta/k} = \pi_{j\Delta} = \pi_\Delta$. For given $\epsilon > 0$, let $\delta_0 = \delta_0(\Delta, \epsilon)$ be as in Lemma 7.3. Since $\{j\Delta/k\}_{j,k\in\mathbb{Z}^+}$ is dense in $\mathbb{R}_+$, we may choose j, k so that $\delta := \Gamma - \frac{j\Delta}{k} \in (0, \delta_0)$. Then

$$\pi_\Delta P^\Gamma = \left(\pi_\Delta P^{\frac{j\Delta}{k}}\right) P^\delta = \pi_\Delta P^\delta,$$

and (7.1) holds by Lemma 7.3. □

Henceforth we will write $\pi = \pi_\Delta$, any $\Delta > 0$.

Theorem 2 follows immediately from Theorem 1 and Proposition 7.2.

Proof of Theorem 3 We first prove (b), and then deduce (a) from (b). Let $\mu_1, \mu_2 \in L_V(\mathbb{R}^N_+)$ be given. For $t \geq 1$, we let $n \in \mathbb{Z}^+$ be such that $t = n + \Delta$ and $\Delta \in [1, 2]$. Then by Proposition 7.1,

$$\begin{aligned}
\left\|\mu_1 P^t - \mu_2 P^t\right\|_V &= \left\|\left(\mu_1 P^\Delta\right) P^n - \left(\mu_2 P^\Delta\right) P^n\right\|_V \\
&\leq C_0 \rho_0^n \cdot \left\|\mu_1 P^\Delta - \mu_2 P^\Delta\right\|_V \\
&\leq C_0 \rho_0^n \cdot \left(C_0 \rho_0 \|\mu_1 - \mu_2\|_V\right).
\end{aligned}$$

For $t < 1$, using the fact that $P^t V \leq V + B \leq (1 + B)V$ where B is as in Proposition 5.1 and $V \geq 1$, we obtain

$$\left\|\mu_1 P^t - \mu_2 P^t\right\|_V = \int_{\mathbb{R}^N_+} P^t V |\mu_1 - \mu_2| \leq (1 + B)\|\mu_1 - \mu_2\|_V.$$

Theorem 3(a) is a special case of part (b), with $\mu_1 = \delta_{\mathbf{x}}$, point mass concentrated at $\mathbf{x}$, and $\mu_2 = \pi$. The expression in (a) follows from

$$\|\delta_{\mathbf{x}} - \pi\|_V \leq \int V(\mathbf{z})(\delta_{\mathbf{x}} + \pi)(d\mathbf{z}) = V(\mathbf{x}) + \|\pi\|_V \leq \left(\|\pi\|_V + 1\right) V(\mathbf{x}).$$ □

7.3 Correlation Decay

Under the conditions of Corollary 4, we have

$$\begin{aligned}
\left|C^\mu_{\xi,\zeta}(t)\right| &= \left|\int \xi(\mathbf{x})\left(\left(P^t\zeta\right)(\mathbf{x}) - \int \left(P^t\zeta\right)(\mathbf{z})\mu(d\mathbf{z})\right)\mu(d\mathbf{x})\right| \\
&\leq \|\xi\| \int \left(\int |\zeta| \left|\delta_{\mathbf{x}} P^t - \mu P^t\right|\right)\mu(d\mathbf{x}) \\
&\leq \|\xi\| \|\zeta\|_V \int \left(\int V \left|\delta_{\mathbf{x}} P^t - \mu P^t\right|\right) \mu(d\mathbf{x}).
\end{aligned}$$

The quantity inside the inner parentheses being, by definition, $\|\delta_{\mathbf{x}} P^t - \mu P^t\|_V$, it follows from Theorem 3 that it is bounded above by $c\rho^t \|\delta_{\mathbf{x}} - \mu\|_V$. We finish by observing that

$$\int \|\delta_{\mathbf{x}} - \mu\|_V \mu(d\mathbf{x}) \le \int \big(V(\mathbf{x}) + \|\mu\|_V\big)\mu(d\mathbf{x}) = 2\|\mu\|_V.$$

Corollary 4 is proved.

References

1. Balint, P., Lin, K.K., Young, L.S.: Ergodicity and energy distributions for some boundary driven integrable Hamiltonian chains. Commun. Math. Phys. **294**(1), 199–228 (2010)
2. Basile, G., Bernardin, C., Olla, S.: Thermal conductivity for a momentum conservative model. Commun. Math. Phys. **287**(1), 67–98 (2009)
3. Bernardin, C., Olla, S.: Fourier's law for a microscopic model of heat conduction. J. Stat. Phys. **121**(3–4), 271–289 (2005)
4. Bonetto, F., Lebowitz, J.L., Rey-Bellet, L.: Fourier's law: A challenge to theorists. Math. Phys. 128–150 (2000)
5. Bricmont, J., Kupiainen, A.: Towards a derivation of Fourier's law for coupled anharmonic oscillators. Commun. Math. Phys. **274**(3), 555–626 (2007)
6. Bunimovich, L., Liverani, C., Pellegrinotti, A., Suhov, Y.: Ergodic systems of balls in a billiard table. Commun. Math. Phys. **146**(2), 357–396 (1992)
7. Carmona, P.: Existence and uniqueness of an invariant measure for a chain of oscillators in contact with two heat baths. Stoch. Process. Appl. **117**(8), 1076–1092 (2007)
8. Cuneo, N., Eckmann, J.-P.: Controlling general polynomial networks. arXiv:1301.1235 (2013, preprint)
9. Derrida, B.: Non-equilibrium steady states: fluctuations and large deviations of the density and of the current. J. Stat. Mech. Theory Exp. **2007**(07), P07023 (2007)
10. Derrida, B., Lebowitz, J.L., Speer, E.R.: Large deviation of the density profile in the steady state of the open symmetric simple exclusion process. J. Stat. Phys. **107**(3), 599–634 (2002)
11. Dolgopyat, D., Liverani, C.: Energy transfer in a fast-slow Hamiltonian system. Commun. Math. Phys. **308**(1), 201–225 (2011)
12. Eckmann, J.P., Young, L.S.: Nonequilibrium energy profiles for a class of 1-d models. Commun. Math. Phys. **262**(1), 237–267 (2006)
13. Eckmann, J.P., Young, L.S.: Rattling and freezing in a 1d transport model. Nonlinearity **24**(1), 207 (2010)
14. Eckmann, J.-P., Pillet, C.-A., Rey-Bellet, L.: Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. Commun. Math. Phys. **201**(3), 657–697 (1999)
15. Gaspard, P., Gilbert, T.: Heat conduction and Fourier's law in a class of many particle dispersing billiards. New J. Phys. **10**(10), 103004 (2008)
16. Grigo, A., Khanin, K., Szasz, D.: Mixing rates of particle systems with energy exchange. Nonlinearity **25**(8), 2349 (2012)
17. Hairer, M., Mattingly, J.C.: Yet another look at Harri's' ergodic theorem for Markov chains. In: Seminar on Stochastic Analysis, Random Fields and Applications VI, pp. 109–117. Springer, Berlin (2011)
18. Harris, T.E.: The existence of stationary measures for certain Markov processes. In: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, vol. 2, pp. 113–124. University of California Press, Berkeley (1956)
19. Khas´minskiĭ, R.Z.: Stochastic Stability of Differential Equations vol. 7. Kluwer Academic, Norwell (1980)
20. Kipnis, C., Marchioro, C., Presutti, E.: Heat flow in an exactly solvable model. J. Stat. Phys. **27**(1), 65–74 (1982)
21. Larralde, H., Leyvraz, F., Mejia-Monasterio, C.: Transport properties of a modified Lorentz gas. J. Stat. Phys. **113**(1), 197–231 (2003)
22. Lefevere, R., Zambotti, L.: Hot scatterers and tracers for the transfer of heat in collisional dynamics. J. Stat. Phys. **139**(4), 686–713 (2010)
23. Lepri, S., Livi, R., Politi, A.: Thermal conduction in classical low-dimensional lattices. Phys. Rep. **377**(1), 1–80 (2003)
24. Lin, K.K., Young, L.S.: Nonequilibrium steady states for certain Hamiltonian models. J. Stat. Phys. **139**(4), 630–657 (2010)

25. Liverani, C., Olla, S., et al.: Toward the Fourier law for a weakly interacting anharmonic crystal. J. Am. Math. Soc. **25**, 555–583 (2011)
26. Luc, R.-B., Thomas, L.E.: Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators. Commun. Math. Phys. **215**(1), 1–24 (2000)
27. Luc, R.-B., Thomas, L.E.: Exponential convergence to non-equilibrium stationary states in classical statistical mechanics. Commun. Math. Phys. **225**(2), 305–329 (2002)
28. Meyn, S.P., Tweedie, R.L., Glynn, P.W.: Markov Chains and Stochastic Stability vol. 2. Cambridge University Press, Cambridge (2009)
29. Prados, A., Lasanta, A., Hurtado, P.I.: Large fluctuations in driven dissipative media. Phys. Rev. Lett. **107**(14), 140601 (2011)
30. Ravishankar, K., Young, L.-S.: Local thermodynamic equilibrium for some stochastic models of Hamiltonian origin. J. Stat. Phys. **128**(3), 641–665 (2007)
31. Rey-Bellet, L.: Nonequilibrium statistical mechanics of open classical systems. In: XIVTH International Congress on Mathematical Physics, pp. 447–454 (2003)
32. Robbins, H.: A remark on Stirling's formula. Am. Math. Mon. 26–29 (1955)
33. Ruelle, D.: A mechanical model for Fourier's law of heat conduction. Commun. Math. Phys. **311**(3), 755–768 (2012)
34. Ryals, B., Young, L.S.: Nonequilibrium steady states of some simple 1-d mechanical chains. J. Stat. Phys. 1–15 (2012)
35. Yarmola, T.: Ergodicity of some open systems with particle-disk interactions. Commun. Math. Phys. **304**(3), 665–688 (2011)